

文章编号: 2095-2163(2022)09-0183-06

中图分类号: TP391

文献标志码: A

# 基于深度学习和视觉语义关系的防振锤滑移目标检测

安金鹏, 王晓春

(新疆信息产业有限责任公司, 乌鲁木齐 830011)

**摘要:** 为了精确定位防振锤滑移异常,保障输电线路安全运行,在输电线路巡检过程中,通常使用视觉检测的方法,对防振锤滑移异常及相关部件进行识别。但输电线路场景复杂多变,常规目标检测算法对防振锤滑移异常检测的精度较差,无法满足实际检测需求。因此,本文提出一种基于深度学习和视觉语义关系的防振锤滑移目标检测方法。该方法依据空间上下文信息,判断所检测目标间的视觉语义关系;联合 Cascade R-CNN 目标检测算法,并利用制定的相应判别规则及约束算法,实现防振锤滑移异常判别。实验结果表明,与常规目标检测算法相比,本文方法对防振锤滑移异常目标识别更加有效,在输电线路巡检中具有较高的理论价值与可观应用前景。

**关键词:** 防振锤; 深度学习; Cascade R-CNN; 视觉语义关系

## Target detection of vibration damper movement based on deep learning and visual semantic relations

AN Jinpeng, WANG Xiaochun

(Xinjiang Information Industry Co., Ltd., Urumqi 830011, China)

**[Abstract]** In order to accurately locate the abnormal sliding of the vibration damper to ensure the safe operation of the transmission line, visual inspection method is usually used to identify the abnormal sliding of the vibration damper and the related components during the inspection process of the transmission line. However, the target detection algorithm has poor accuracy in the detection of vibration damper slip anomalies in the experiment and cannot meet the actual detection needs. Therefore, in this paper, a vibration damper slip target detection method based on deep learning and visual semantic relationship is proposed. This method judges the visual semantic relationship between the detected targets according to the spatial context information. It combines the Cascade R-CNN target detection algorithm and uses the corresponding discrimination rules and constraint algorithms developed to realize the vibration damper slip anomaly discrimination. The experimental results show that compared with the conventional target detection algorithm, the method proposed in this paper is more effective for the abnormal target recognition of vibration damper slippage, and has higher theoretical value and broad application prospects in the inspection of transmission lines.

**[Key words]** vibration damper; deep learning; Cascade R-CNN; the visual semantic relationship

## 0 引言

长期稳定供电能够保障国家经济的发展以及居民的稳定生活,因此输电线路的安全防护至关重要。国内外常见的电力导线断股原因,主要是微风振动导致的线路疲劳,最终形成断股。而防振锤的安放位置在保护电力导线中起着关键作用,因此对防振锤是否存在滑移异常检测就显得尤为重要。传统人工筛检方式,对于防振锤滑移目标的检测成本高、效率低,并对大量检测数据不能做出及时的处理。针对于此,国内外一些学者开始着手研究相关的机器视觉识别方法。判断防振锤是否滑移,首先需要能够识别出防振锤部件。关于防振锤部件的识别在国内外已有大量的研究资料。如:文献[1]中提出对

图像的LBP特征、Haar特征及HOG特征进行组合,来识别防振锤所在位置。经验证,精确度能够满足检测要求。文献[2]提出利用改进的Canny算法,对采集到的防振锤螺栓图像进行边缘提取后,利用改进的Hough变换,对防振锤紧固螺栓的框架和中心进行精确定位,从而完成防振锤部件的位置判断,实验结果表明,此方法非常可靠有效。文献[3]中提出了一种基于深度学习的多特征融合检测方法,该方法首先利用卷积神经网络对防振锤进行训练,随后使用多个特征融合后的特征图像识别防振锤。经实验证明,此方法不仅能够提高防振锤检测精度,同时还提升了检测速率。

综上所述,虽然大量资料表明对于防振锤等相关部件的检测取得了突破性的进展,但对于防振锤

**作者简介:** 安金鹏(1983-),男,硕士,高级工程师,主要研究方向:信息通信、人工智能;王晓春(1978-),女,学士,高级工程师,主要研究方向:目标识别、信息通讯。

收稿日期: 2022-03-01

滑移异常诊断还未有后续研究。

本文结合前人经验,提出一种基于深度学习和视觉语义关系的防振锤滑移异常目标检测方法。首先基于 Cascade R-CNN 网络<sup>[4]</sup>训练了多级联的检测器,实现对防振锤部件及其相关联部件的检测与定位。依据空间上下文信息,判断所检测目标间的语义关系。基于输电线路相关规范,制定以下判断规则:

(1)防振锤与其它部件或两防振锤间相对距离小于规范时,判断为防振锤滑移。

(2)当图像中出现  $n(n \geq 2)$  个防振锤时,利用几何约束关系及聚类分析算法实现异常判别。

最终依据上述制定的相应视觉语义判别规则及约束算法,结合 Cascade R-CNN 目标检测方法,实现防振锤滑移异常判别。

## 1 防振锤滑移目标检测算法

### 1.1 基于 Cascade R-CNN 的相关部件目标识别与定位

#### 1.1.1 Cascade R-CNN 简介

CascadeR-CNN 是目前两阶段目标检测算法中表现较为优异的网络,算法中提出了一种多阶段目标检测体系结构,即级联 R-CNN<sup>[5]</sup>。级联网络由一系列经过提高 *IoU* 阈值训练的检测器组成,检测器是按顺序训练的,每个阶段负责预测特定类别下的置信度分数给定的前景/背景阈值。由于在目标检测时需要为每个目标实例分配一个类别标签(如防振锤、线夹等),因此使用检测器的输出作为下一个检测器的训练集,这种方式能够逐步提高假设质量,保证了所有检测器的正训练集大小相等,并使过度拟合达到最小化。

#### 1.1.2 多尺度骨干网络 Res2Net

卷积网络中的骨干网络主要就是用于提取图像不同通道中不同颜色的语义特征。例如,ResNet-50、ResNet-101<sup>[6]</sup>、ResNeXt-101<sup>[7]</sup>。目前常用的骨干网络特征提取大多采用分层方式表示多尺度特征,而 Res2Net 则在更细粒度的级别上提高了多尺度表示能力。区别于利用不同分辨率的特征来提高多尺度的能力<sup>[8-10]</sup>,Res2Net 在更细粒度的级别上提供了多尺度表示能力。为了实现这个目标,Res2Net 用一组较小的滤波器组替换了  $n$  个通道的  $3 \times 3$  滤波器,每个滤波器组都有  $w$  个通道。

基本卷积模块与 Res2Net 模块之间的差异如图 1 所示。图 1 中,“\*”代表卷积操作。经过  $1 \times 1$  卷

积后,将特征图平均分成  $s$  个用  $x_i$  表示的特征子集。与输入特征图相比,每个特征子集  $x_i$  具有相同的空间大小,但通道数为  $1/s$ 。除  $x_1$  外,每个  $x_i$  都对应一个的  $3 \times 3$  卷积。此时,  $y_i$  为特征子集  $x_i$  与  $K_{i-1}()$  的输出相加,并与输入  $K_i()$  作用后的卷积结果。为了在增加  $s$  的同时减少参数,网络中省略了  $x_1$  的  $3 \times 3$  卷积。因此,  $y_i$  可以表示为:

$$y_i = \begin{cases} x_i & i = 1 \\ K_i(x_i) & i = 2 \\ K_i(x_i + y_{i-1}) & 2 < i \leq s \end{cases} \quad (1)$$

其中,  $i \in \{1, 2, \dots, s\}$ ,  $K_i()$  表示  $3 \times 3$  卷积结果。

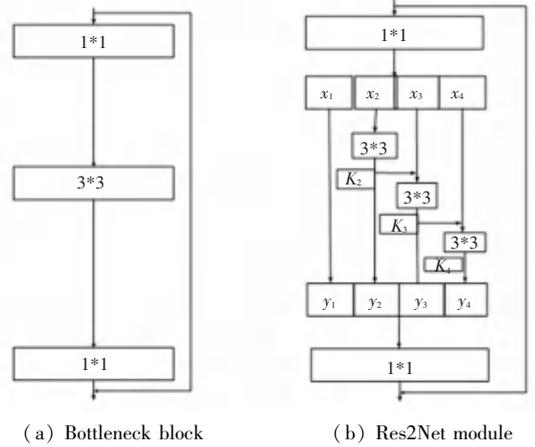


图 1 基本卷积模块与 Res2Net 模块之间对比

Fig. 1 Comparison between the basic convolutional module and the Res2Net module

这些较小的过滤器组以类似分层残差的方式连接,用来增加输出特征可以表示的尺度数量。具体来说,将输入特征图分为几组,过滤器首先从一组输入特征图中提取特征,然后将前一组的输出特征与另一组的输入特征图一起发送到下一组过滤器<sup>[11]</sup>。重复上述过程,直到处理了所有输入特征图后为止。最后将来自所有组的特征图融合,并发送到另一组  $1 \times 1$  过滤器,以完全融合信息。所输入的信息会经过不同的特征提取通道,每当其通过  $3 \times 3$  过滤器时,等效感受野<sup>[12]</sup>就会增加。这种先拆分、后融合的策略能够使卷积更高效地处理特征,使网络的特征提取能力更强大。

#### 1.1.3 整体网络结构

目标检测架构如图 2 所示。首先,将来自无人机巡检时拍照的输电线路检测图片输入到 Res2Net 网络进行目标特征提取;此外,在网络 NECK 结构中还引入了 FPN 结构,在基本不增加原有模型计算量的情况下,大幅度提升了较小缺陷目标检测的性能,

FPN 结构通过自上而下的路径和横向连接,将低分辨率语义的强特征与高分辨率语义的弱特征相结合,构建了具有更多层的特征金字塔信息,能够获取所检测目标准确定位的低级信息和高级信息,使得

特征更具有鲁棒性,可有效提高较小目标的检测精度,所构建的目标检测网络能够极大地提升防振锤相关目标检测精度,确保后续防振锤滑移目标判断的正确性。

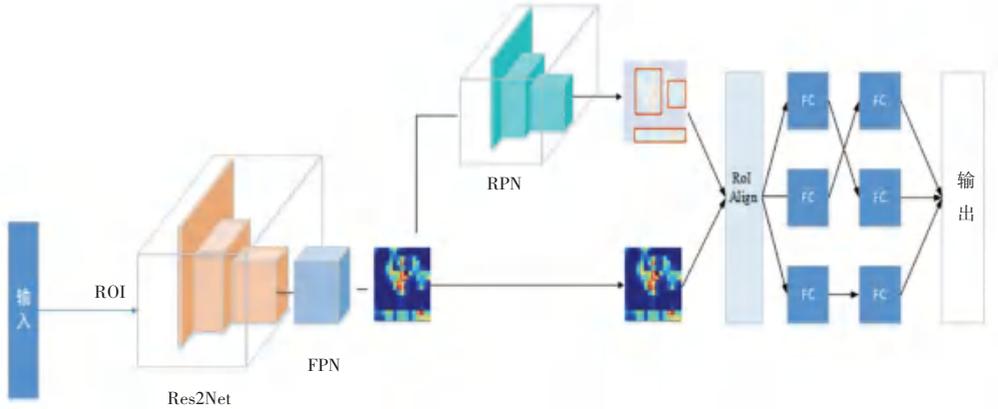


图2 目标检测网络结构图

Fig. 2 Target detection network structure diagram

### 1.2 防振锤与相关部件视觉关系

视觉关系是图像中目标间的相互作用,存在于由多个物体组成的图片中,而每个目标间的关系在语义上是相互关联的。当分析一幅图片时,视觉关系往往会作为已知的知识或经验参与到分析过程中,帮助人们辨别事物。针对本文提出的问题,简单来说,当需要判断防振锤目标是否滑移时,只需要运用相关目标之间的关系,将目标所处的环境与姿态描述出来,说明物体之间语义关系,并将其转化为特

征信息输入至神经网络中即可。

图3展示了防振锤可能出现滑移的多种姿态与位置变化情况。分析可知,图3(a)中存在多线电缆时,会出现防振锤并排现象;图3(b)中,防振锤与绝缘子相对距离超过标准安装距离(或接触),可判断为防振锤滑移;图3(c)中,线夹两侧防振锤相对距离超过标准安装距离(或接触),可判断为防振锤滑移;图3(d)中,防振锤与连接金具相对距离超过标准安装距离(或接触),可判断为防振锤滑移。



(a) 多条线路防振锤并排

(b) 防振锤靠近绝缘子

(c) 防振锤靠近线夹

(d) 防振锤靠近连接金具

图3 防振锤位置与姿态

Fig. 3 The position and attitude of the vibration damper

针对上述情况,具体计算方法可表示为:首先根据神经网络前向反馈,得出防振锤及其相关联部位图像的坐标,并计算出两者在图像中的距离 $d$ ;通过相机参数,推算两者实际距离 $D$ ;根据经验值设置滑动阈值 $T$ ,若 $|D - D_r| \leq T$ 可断定为防振锤不存在滑移,若 $|D - D_r| \geq T$ 可断定为防振锤滑移。这里, $D_r$ 为标准安装距离。计算公式如下所示:

$$d = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \quad (2)$$

$$D = f(d) \quad (3)$$

其中, $(x_1, y_1)$ 为防振锤的图像坐标; $(x_2, y_2)$ 为防振锤相邻部位的图像坐标; $f(d)$ 为图像像素距离与空间距离变换的求解公式<sup>[13]</sup>。

结合模型预测结果,本文利用位置关系描述防振锤与相关目标(绝缘子、连接金具、线夹)的空间关系,距离关系可直接判断防振锤是否出现滑移异常。

## 2 实验及分析

### 2.1 防振锤异常及相关部件的常规目标检测方法

实验训练和评估模型所使用的硬件环境为Tesla V100-DGX Work Station,其中搭载了4块NVIDIA Tesla V100 Tensor Core GPU。考虑到可编程性和灵活性,选择Pytorch作为深度学习框架<sup>[14]</sup>,并使用经过ImageNet-1000预训练的模型,在输电线路数据集上做迁移学习,随机初始化其它层参数。使用Cascade-RCNN级联网络,训练防振锤、线夹、连接金具、绝缘子四类目标。训练过程中,判断正负样本的IoU阈值分别设为0.5、0.6、0.7,输入图像训练尺寸为((1 200, 600), (2 400, 1 200)),学习率初始化为0.01。

本文所使用训练集共有22 873张图片,经过150 000轮迭代,选取最优模型在测试集上进行测试。测试分为2部分:

(1)使用7 600张测试集数据对绝缘子、线夹等相关部件进行测试。

(2)收集2 831张含有防振锤滑移异常与无异常的数据,验证常规方式下Cascade-RCNN级联网络对防振锤滑移异常的检测效果。

采用目标识别应用中常用的评价标准:精准率(precision)和召回率(recall),评估算法的有效性<sup>[15-16]</sup>。推导得到的公式分别如下:

$$Precision = TP / (TP + FP) \quad (4)$$

$$Recall = TP / (TP + FN) \quad (5)$$

其中,TP表示防振锤滑移目标被算法正确检测的数量,FP表示被算法错误分类为防振锤滑移目

标的数量;FN表示算法遗漏的防振锤滑移目标。

测试结果详见表1。针对绝缘子、线夹等部件目标识别的测试结果显示,Cascade-RCNN级联网络检测效果显著,能够实现0.9以上的平均检测精度,得到的召回率均在0.95以上,说明部件目标漏检情况极少发生。然而,在对2 831张混合有防振锤滑移异常与无异常的数据进行验证时,常规方式下Cascade-RCNN级联网络对防振锤滑移异常的检测效果只能达到0.584的平均检测精度,说明单纯利用Cascade-RCNN级联网络进行目标识别,网络对目标的形状、轮廓、纹理、颜色等特征提取效果显著,而对目标相对位置变化并不敏感,无法满足输电线路防振锤滑移异常的检测需求。但Cascade-RCNN级联网络对相关部件识别效果甚佳,为后期的防振锤滑移识别工作提供了良好的基础。

表1 常规目标检测精度结果

Tab. 1 Model test results on common components

类别标签	召回率	准确率	AP
绝缘子	0.983	0.875	0.921
线夹	0.968	0.766	0.895
连接金具	0.951	0.810	0.876
防振锤	0.988	0.881	0.937
防振锤滑移异常	0.631	0.561	0.584

### 2.2 基于视觉语义关系的防振锤滑移识别

基于目标在语义上的相互关联性,利用本文提出的目标检测方法,通过最优模型判断目标,并结合语义位置关系、距离关系,判断在图片档距范围内是否出现防振锤滑移异常。首先将部件识别结果及防振锤相关部件空间位置关系映射到向量空间,并通过距离约束,判定防振锤滑移异常。判断过程如下:

(1)判断存在防振锤目标的个数。

(2)判断是否出现防振锤相关联部件目标(如绝缘子、线夹、连接金具等)。

(3)当图像中出现 $n(n \geq 2)$ 个防振锤时,需利用几何约束关系及聚类分析算法实现异常判别。将图片中展示的二维空间,通过Cayley-Menger行列式提供的距离几何约束,结合加权质心计算约束距离,进行判断。

(4)若同时出现防振锤相关联部件目标,还需计算出相邻部件与防振锤之间的 $x$ 轴间距,并转化为世界坐标,通过距离约束判定防振锤滑移异常。

基于此,又对研发出的算法流程做具体描述如下:

算法 判断防振锤滑移

```

输入  $Detection\_results D = \{box, name, score\}_{i=1}^k$ ,
防振锤数  $n$ , 标准安装距离  $D_r$ , 滑动阈值  $T$ 
if  $n = 0$  then 无防振锤滑移目标
else
  if  $n \geq 2$ 
    if  $Cayley - Menger(box) > T$  then return 存在
    防振锤滑移,  $(box, score)$ 
    else 输出无防振锤滑移
  if 绝缘子 in  $D$  or 线夹 in  $D$  or 连接金具 in  $D$  then
     $x_i, y_i = box[j][0], box[j][1]$ 
     $x_j, y_j = box[j][0], box[j][1]$ 
     $d = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}$ 
     $D = f(d)$ 
    if  $|D - D_r| \geq T$  then return 存在防振锤滑移,
     $(box, score)$ 
    else 输出无防振锤滑移
end

```

同样使用 2 831 张混合有防振锤滑移异常与无异常的数据, 对上述算法步骤进行验证。在实验中,

拟将检测所提算法的有效性, 并对检测结果进行了可视化分析, 测试结果如图 4 所示。由图 4 可见, 虽然防振锤滑移目标的形态位置变化较大、目标大小不一, 增加了检测复杂度, 但加入视觉语义关系的检测算法, 检测到防振锤滑移异常正确率却非常高。算法精度测试结果见表 2。表 2 中数据显示, 基于视觉语义关系的目标检测算法在防振锤滑移异常中的表现很大程度上优于 Cascade R-CNN 级联网络。与其相比, 基于视觉语义关系的防振锤滑移异常目标检测, 在召回率方面提升了 38.4%, 准确度方面提升了 30.3%。基于视觉语义关系目标检测方法的结果如图 4 所示。分析可知, 图 4(a)与图 4(c)中, 为并排线路中防振锤出现滑移现象; 图 4(b)中显示防振锤滑移状况非常严重, 而且出现了碰撞现象; 图 4(d)中, 防振锤距离线夹部件小于标准安装距离, 也被判别为防振锤滑移。综合上述可视化结果, 本文所提算法在不同场景下对防振锤滑移异常检测都得到不错的效果。该模型能够应对多种环境变化, 满足无人机巡检的实际需求, 具有很好的鲁棒性和泛化能力。



图 4 防振锤滑移目标检测结果

Fig. 4 Detection results of vibration damper sliding target

表 2 算法精度测试结果

Tab. 2 Results of different algorithms

防振锤滑移	召回率	准确率	AP
Cascade R-CNN 算法	0.631	0.561	0.584
基于视觉语义关系的目标检测方法	0.874	0.730	0.789

注: 数值表示各精度, 值为 1 时精度最高, 且无单位。

### 3 结束语

针对无人机输电线防振锤滑移异常对线路带来的危害, 本文结合深度学习和视觉关系目标检测, 提出了一种基于 Cascade R-CNN 算法联合视觉语义

关系进行防振锤滑移异常检测算法,能够对输电线路防振锤滑移异常进行自动识别定位。经巡检图像实践测试,所提方法实现了防振锤滑移异常检测且效果显著,由此得到以下结论:

(1)本文提出的输电线路中防振锤滑移目标的检测方法,依据空间上下文信息,判断所检测目标间的视觉语义关系,并联合 Cascade R-CNN 目标检测算法及约束算法,实现了防振锤滑移异常判别。最终实验结果表明,本文所提的算法与常规目标检测算法相比在召回率方面提升了 38.4%,准确度提升了 30.3%,平均准确度为 78.9%,能够有效地检测出 3 类防振锤滑移异常目标。

(2)对实验结果进行了不同场景的测试与分析,所训练模型能够应对多种环境变化,满足无人机巡检的实际需求,具有很好的鲁棒性和泛化能力。

(3)本文算法能够及时监测输电线路中防振锤滑移所引起的线路隐患,完成输电线路缺陷自主巡检任务,有效地避免了输电线中防振锤滑移造成的线路问题,具有较高的理论价值和实际应用前景。

## 参考文献

[1] 金立军, 闫书佳, 刘源. 基于类 Haar 特征与级联 AdaBoost 算法的防震锤识别[J]. 系统仿真学报, 2012, 24(09):1806-1809.

[2] LIU Yanmei, WEN Shihua, CHEN Zhen, et al. Research of the Anti-vibration Hammer Resetting Robot Based on Machine Vision [C]//第 32 届中国控制与决策会议. 合肥:《控制与决策》编辑部, 2020:627-631.

[3] 洪刚. 基于深度学习的无人机图像防震锤检测研究[J]. 内蒙古科技与经济, 2019(20):70-71.

[4] CAI Zhaowen, VASCONCELOS N. Cascade R-CNN: Delving into high quality object detection [C]// 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR).

Salt Lake City: IEEE, 2018:1-9.

[5] AHMED M N, TOOR A S. Detection of objects in images using region-based convolutional neural networks: USA, 9, 996, 772 [P]. 2018-06-12.

[6] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition [C]// 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas: IEEE, 2016:770-778.

[7] XIE Saining, GIRSHICK R, DOLLÁR P, et al. Aggregated residual transformations for deep neural networks [C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, HI, USA:IEEE, 2017:5987-5995.

[8] CHEN C F, FAN Q, MALLINAR N, et al. Big-little net: An efficient multi-scale feature representation for visual and speech recognition [J]. arXiv preprint arXiv:1807.03848, 2018.

[9] CHEN Yunpeng, FAN Haoqi, XU Bing, et al. Drop an octave: Reducing spatial redundancy in convolutional neural networks with octave convolution [C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. Seoul, Korea (South):IEEE, 2019: 3435-3444.

[10] CHENG B, XIAO R, WANG J, et al. High frequency residual learning for multi-scale image classification [J]. arXiv preprint arXiv:1905.02649, 2019.

[11] 乔玲玲. 图像分割算法研究及实现 [D]. 武汉:武汉理工大学, 2009.

[12] LIM B, SON S, KIM H, et al. Enhanced deep residual networks for single image super-resolution [J]. arXiv preprint arXiv:1707.02921, 2017.

[13] 张进秋, 李响. 图像像素距离与空间距离变换公式的求解 [J]. 软件导刊, 2009, 8(12):179-180.

[14] KETKAR N, MOOLAYIL J. Introduction to pytorch [M]//Deep learning with python. Berkeley, CA: Apress, 2017: 195-208.

[15] VICENTE S, CARREIRA J, AGAPITO L, et al. Reconstructing PASCAL VOC [C]//2014 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Columbus, OH, USA: IEEE, 2014:41-48.

[16] LONG Xitian, ZHENG Zhe, LIU Rui, et al. Cascaded feature enhancement network model for real-time video monitoring of power system [J]. Energy Reports, 2021, 7:8485-8492.

(上接第 182 页)

[10] LYU C, HUO Ziqiang, CHENG Xin, et al. Distributed optical fiber sensing intrusion pattern recognition based on GAF and CNN [J]. Journal of Lightwave Technology, 2020, 38(15): 4174-4182.

[11] 吴虎, 孔勇, 王振伟, 等. 基于端点检测与信号重组的光纤分布式传感信号识别 [J]. 光子学报, 2021, 50(11):123-130.

[12] 李祥兵, 肖合林. 基于 Matlab 的小波阈值折衷去噪算法研究 [J]. 电脑开发与应用, 2009, 22(06):4-6.

[13] 黄洋, 赵风海, 卢景. 语音信号处理中双门限端点检测算法的改进 [J]. 南开大学学报(自然科学版), 2021, 54(02):58-62.

[14] 杨维民, 安征. 结合图像分割和 ResNet18 的危险驾驶行为识别 [J]. 信息通信, 2020(06):7-8.

[15] WANG Yanping, GONG Dianjun, PANG Liping, et al. RVFL-based optical fiber intrusion signal recognition with multi-level wavelet decomposition as feature [J]. Photonic Sensors, 2018, 8(3): 234-241.